Boolean-Valued Semantics for the Stochastic λ-Calculus

Giorgio Bacci
Aalborg University, Denmark

Robert Furber
Aalborg University, Denmark

Dexter Kozen
Cornell University, USA

Radu Mardare
Aalborg University, Denmark

Prakash Panangaden
McGill University, Canada

Dana Scott
Carnegie Mellon University, USA

Abstract

The ordinary untyped λ-calculus has a set-theoretic model proposed in two related forms by Scott and Plotkin in the 1970s. Recently Scott showed how to introduce probability by extending these models with random variables. However, to reason about correctness and to add further features, it is useful to reinterpret the construction in a higher-order Boolean-valued model involving a measure algebra. We develop the semantics of an extended stochastic λ-calculus suitable for modeling a simple higher-order probabilistic programming language. We exhibit a number of key equations satisfied by the terms of our language. The terms are interpreted using a continuation-style semantics with an additional argument, an infinite sequence of coin tosses, which serves as a source of randomness. We also introduce a fixpoint operator as a new syntactic construct, as β-reduction turns out not to be sound for unrestricted terms. Finally, we develop a new notion of equality between terms interpreted in a measure algebra, allowing one to reason about terms that may not be equal almost everywhere. This provides a new framework and reasoning principles for probabilistic programs and their higher-order properties.

Keywords Stochastic λ-calculus, Boolean-valued models, random variables, denotational semantics.

1 Introduction

Probabilistic programming languages [4–7, 11–13, 21] have become popular recently, sparked by renewed interest in verification and machine learning. The subject began with work by Saheb-Djaromi [16] on a probabilistic version of LCF and, in an imperative first-order language by [11]. There has been significant recent interest in extending to higher-order functional languages [7, 12, 13, 20]. The higher-order functional paradigm allows one to integrate probability distributions smoothly into the programming language through the probability monad, but finding a cartesian-closed category [9] that can incorporate higher-order features as well as appropriate probabilistic constructions [8] has proven elusive. Only recently [7] has a suitable category been constructed that satisfies all desiderata.

LICS '18, July 9–12, 2018, Oxford, United Kingdom

ACM ISBN 978-1-4503-5583-4/18/07…$15.00
https://doi.org/10.1145/3209108.3209175

In the present paper, we take an entirely new approach to the semantics of higher-order probabilistic computation. In [18] Scott, one of the authors of the present paper, proposed a way of incorporating random variables into a certain kind of model of the untyped λ-calculus by using the continuity of the λ-calculus operations modeled by enumeration operators on the powerset of the integers. This set-theoretic model suggests at once incorporating higher types, but to do this requires a nonstandard Boolean-valued interpretation of set theory. Boolean-valued models (see [1] for history and a basic exposition) were employed by Scott [17] to construct models of set theory in order to obtain independence results. The independence of the Continuum Hypothesis was obtained by introducing an arbitrarily large set of real-valued random variables. In this model, the random variables turned out to be the real numbers of the Boolean-valued logic. The continuity of real algebra has an analogue in the continuity of operations in the powerset model. The measure algebra of a standard Borel space, a complete Boolean algebra, is used to bring this idea to fruition. Ordinary logical propositions take elements in this Boolean algebra as truth values, instead of in the simple two-element Boolean algebra.

The ideas behind the present work specify and extend the basic intuitions briefly outlined in [18]. The primary goal here is to develop an equational theory in which equations between stochastic λ-terms have probabilistic meaning and are valued in a complete Boolean algebra. The intention is to provide reasoning principles for evaluating the equality of λ-terms under various program transformations.

The notions of equality and invariance are subtle in the presence of probabilities. In the calculus described below, there is a probabilistic choice operator $\oplus$, which captures the idea that a choice is to be made between two terms based on a random process. The source of randomness is called a *tossing process*: a process that generates a sequence of fair coin tosses, the outcomes of which are used to resolve the probabilistic choices. In general, equality of terms does not mean that identical values are produced, as the final values will depend on the tosses. Instead, we interpret equality statements as elements of a measure algebra formed from the usual measurable sets quotiented by the ideal of negligible sets. Given a tossing process $\mathcal{T}$, a pair of closed terms M, N will define a set of tossing sequences where they agree $[\![M = N]\!]_{\mathcal{T}}$, which is an element of the Boolean algebra. This may be the top element—corresponding to certainty—or something else. However, we would like statements not to be dependent on the specific outcome of a tossing process; rather, we would prefer that truth values of equations be invariant under certain changes in the tossing process. Accordingly, we define a relation $\approx$ on the elements of the Boolean algebra to capture the idea that two truth values of an equation, say $[\![M = N]\!]_{\mathcal{T}}$ and $[\![M = N]\!]_{\mathcal{T}'}$ for different tossing processes $\mathcal{T}$ and $\mathcal{T}'$, are related by an automorphism of the Boolean algebra. We write $[\![M = N]\!]_{\mathcal{T}} \approx$

$[\![M = N]\!]_{\mathcal{T}'}$ when this occurs. Many of the equalities that we establish are stated in this way, and the automorphisms relating them are constructed.

A second subtlety is that we often prove results of the form $[\![M = K]\!]_{\mathcal{T}} \approx [\![N = K]\!]_{\mathcal{T}}$, for closed stochastic terms M, N and closed term K of the classic untyped λ-calculus (here called a stable term), instead of proving for example that $[\![M = N]\!]_{\mathcal{T}}$ evaluates to the top element of the Boolean algebra. Here we are using the idea that a tossing process, once it has resolved the choices, makes a term of the stochastic λ-calculus look like an ordinary λ-term. In these cases one cannot prove that $[\![M = N]\!]_{\mathcal{T}}$ is the top element directly, since this might not be true, but the weaker statement above serves to replace this statement.

Our main contributions are:

1. We develop a use of random variables in this framework in order to identify a class of tossing processes that can serve as a source of randomness in probabilistic programs. We show that with this choice, the semantics is invariant under automorphisms of the measure algebra effected by remapping the tossing process. This provides a canonical meaning to programs.

2. We introduce the stochastic λ-calculus by augmenting the ordinary λ-calculus with a probabilistic choice construct. We flesh out the continuation-passing semantics proposed in [18], with the crucial new observation that β-reduction is not sound for all terms with probabilistic choice. To compensate, we need to introduce an explicit fixpoint operator in order to have recursive programs.

3. We develop a Boolean-valued reasoning framework for the stochastic λ-calculus and prove soundness results with respect to the continuation-passing semantics.

The main technical contributions and novelty are in items 1 and 3. In order to obtain the invariance of the semantics, it was necessary to identify a rather subtle condition that we call *monolithic*. This is the part that required the deepest foray into the technicalities of measure theory. As far as we know, Item 3 is a completely new way of thinking about equational logic. We have only developed the rudiments here for the purposes of the present investigation, but there is clearly a much deeper theory to be explored.

We have not developed an operational semantics or rewrite rules, but have left these investigations for future work. However, because of the restrictions on β-reduction, all probabilistic choices for a term in the argument position must be resolved before applying the function; thus it most resembles a call-by-value strategy, but of course one cannot talk about evaluation strategies in the absence of a reduction system.

2 Standard Probability Spaces

In this section we introduce a few concepts and results regarding standard probability spaces. The concepts of disintegration of a space and of monolithic maps between spaces are essential.

Definition 2.1. Given a measurable space (X, Σ) and a probability measure μ on it, (X, Σ, μ) is a *standard probability space*[1] iff (X, Σ) is Borel isomorphic to a Polish space[2] equipped with its Borel algebra[3].

Consider, e.g., the set $2^{\mathbb{N}}$ of infinite binary sequences with the Cantor topology[4], which has, as basic open sets, the sets $\{\alpha \mid x \prec \alpha\}$, where α ranges over $2^{\mathbb{N}}$, $x \in \{0,1\}^*$, and $\prec$ denotes prefix. Let $\mathcal{B}$ be the Borel σ-algebra of the Cantor topology.

The (fair) *coin-flipping* probability measure[5] P on $\mathcal{B}$ is generated by its values on intervals:

$$P(\{\alpha \mid x \prec \alpha\}) = 2^{-|x|}.$$

The measure space $\Omega = (2^{\mathbb{N}}, \mathcal{B}, P)$ is a standard probability space that we will use in the rest of this paper.

We consider measure-preserving maps between standard probability spaces. The category we use, **Meas**/0 has maps identified if they are equal almost everywhere, *i.e.* except for on a null set.

Definition 2.2. If $f_j : (X_j, \Sigma_j, \mu_j) \to (Z, \Xi, \xi)$, $j = 1, 2$ are two measure-preserving maps with common codomain, we say $f_1 \cong f_2$ or $(X_1, f_1) \cong (X_2, f_2)$ if there exists a measure-preserving isomorphism $i : X_1 \to X_2$ such that $f_2 \circ i = f_1$, except on a subset of X_1 of measure 0.

There is a measure-preserving Borel isomorphism between any standard probability space (X, Σ, μ) whenever μ is *atomless* (*i.e.* all singletons have measure 0), and Ω.

Definition 2.3. Let $f : (X, \Sigma, \mu) \to (Y, \Theta, \nu)$ be a measure- preserving map between standard probability spaces. A set $S \in \Sigma$ is *1-sheeted with respect to* f if for all $y \in Y$ we have that $S \cap f^{-1}(y)$ has at most 1 element. A map f is *monolithic* if it has no 1-sheeted sets of positive measure.

Note that S is 1-sheeted if the restriction of f to S is injective. For instance, $2^{\mathbb{N}}$ is a 1-sheeted set of measure 1 for the identity function $\mathrm{id} : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$, and the set $\{(a_n) \in 2^{\mathbb{N}} \mid a_1 = 0\}$ is a 1-sheeted set of measure $\frac{1}{2}$ for the function $\mathbf{tail} : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$, which takes all but the first element of a sequence to return a sequence.
Next, f is monolithic whenever all such one-sheeted measurable sets have measure 0. For example, the map $\mathbf{evens} : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ which takes every second member of a sequence to construct a sequence *is* monolithic (proven in Section 4), while id and **tail** are not.

The following theorem provides a useful characterization of the concept of monolithic map[6]

Theorem 2.4. *Let $f : (X, \Sigma, \mu) \to (Y, \Theta, \nu)$ be a measure-preserving map of standard probability spaces. Then f is monolithic iff there exists a standard probability space (Z, Ξ, ξ) such that all points of Z have ξ-measure 0, and $(Z \times Y, \pi_2)$ is isomorphic to (X, f); where π_2 denotes the second projection.*

We conclude this section with a useful result regarding the composition of monolithic maps.

[1] Our concept of standard probability space generalizes Rokhlin's original one in the sense that we do not insist on μ being Lebesgue-complete.

[2] A *Polish space* is the topological space induced by a complete separable metric space.
[3] The *Borel algebra* of a topology is the σ-algebra generated by the open sets.
[4] The space is the topological power of ω copies of the discrete space $2 = \{0, 1\}$.
[5] This is the Haar measure on $2^{\mathbb{N}}$ as a compact group based on $mod\ 2$ addition.
[6] This concept is formulated in terms of *decompositions* in [15, §3.1], which are better known as *disintegrations*[2, §452 E].

Lemma 2.5. *Let $f : (X, \Sigma, \mu) \to (Y, \Theta, \nu)$ and $g : (Y, \Theta, \nu) \to (Z, \Xi, \xi)$ be measure-preserving maps. If $S \in \Sigma$ is 1-sheeted with respect to $g \circ f$, then it is 1-sheeted with respect to f. Therefore if f is monolithic, then $g \circ f$ is monolithic.*

Proof. Let $S \in \Sigma$ be a 1-sheeted set with respect to $g \circ f$, *i.e.* for all $z \in Z$, $S \cap f^{-1}(g^{-1}(z))$ has cardinality at most 1. If $y \in Y$, then

$$S \cap f^{-1}(y) \subseteq S \cap f^{-1}(g^{-1}(g(y)))$$

so $f^{-1}(y) \cap S$ has cardinality at most 1, so S is 1-sheeted with respect to f.

The statement about 1-sheeted sets of positive measure then follows by taking the contrapositive. □

3 Topology and measure of $\mathcal{P}(\mathbb{N})$

Let $\mathbb{N}$ be the set of *natural numbers*, $\mathcal{P}(\mathbb{N})$ its powerset and $\mathcal{P}_{fin}(\mathbb{N})$ the set of finite subsets of $\mathbb{N}$. We follow [14, 18] to identify some structure on $\mathbb{N}$ that is relevant to our purpose.

The *pairing function* puts the set of pairs of natural numbers into a one-to-one correspondence with the positive integers by: $(m, n) = 2^n(2m+1)$. The function $\langle\,\rangle$ puts finite sequences of natural numbers into a one-to-one correspondence with $\mathbb{N}$ by nested pairing and this can be used to enumerate finite subsets of $\mathbb{N}$ using a function that we call *set*. Now we can define the *Kleene star* for $X \subseteq \mathbb{N}$:

$$X^* = \{n \mid set(n) \subseteq X\}.$$

All these identify different structures on $\mathbb{N}$ giving us

$$\mathbb{N} \cong 1 + (\mathbb{N} \times \mathbb{N}) \cong \mathbb{N}^*.$$

Enumeration operators are identified with sets $F \subseteq \mathbb{N}$ which operate on $X \subseteq \mathcal{P}(\mathbb{N})$ through the binary operation of *application*:

$$F(X) = \{m \in \mathbb{N} \mid \exists s \in X^*.(s, m) \in F\}.$$

The intuition here is that while enumerating the elements in X, one can also enumerate the elements of X^* and the pairs in F: every match between a sequence number $s \in X^*$ and the first term of a pair $(s, m) \in F$ witnesses the fact that $m \in F(X)$. Following [3, 18], we say that a set $A \in \mathcal{P}(\mathbb{N})$ is *enumeration reducible* to a set $B \in \mathcal{P}(\mathbb{N})$ when there exists a *recursively enumerable* set $F \in \mathcal{P}(\mathbb{N})$ such that $A = F(B)$. Hence, the *computable* enumeration operators are those given by recursively enumerable sets F.

3.1 The positive topology on $\mathcal{P}(\mathbb{N})$

The *positive topology* on $\mathcal{P}(\mathbb{N})$ is induced by the sets

$$Q_n = \{X \mid n \in X^*\}.$$

The continuity of a function $\Phi : \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\mathbb{N})$ in the positive topology can be characterized as follows:

$$m \in \Phi(X) \text{ iff } \exists n \in X^* \text{ s.t. } m \in \Phi(set(n)).$$

Viewing $\mathcal{P}(\mathbb{N})$ as an algebraic lattice, the positive topology coincides with the Scott topology on $\mathcal{P}(\mathbb{N})$, which has basic open sets $\{b \mid a \subseteq b\}$, where $a \subseteq \mathbb{N}$ is finite; and positive continuity coincides with Scott continuity.

The following two results are proven in [18].

Theorem 3.1. *The application operation $F(X)$ is continuous as a function of two variables on $\mathcal{P}(\mathbb{N})$.*

Theorem 3.2. *For every continuous function $\Phi : \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\mathbb{N})$ there is a largest set F such that for any $X \in \mathcal{P}(\mathbb{N})$, $\Phi(X) = F(X)$, where $\Phi(X)$ denotes ordinary function application, while $F(X)$ is application in the set-based model. In fact, F can be directly defined by $F = \{0\} \cup \{(n, m) \mid m \in \Phi(set(n))\}$.*

In view of this fact, we define λ-abstraction on $\mathcal{P}(\mathbb{N})$ as follows:

$$\lambda X.F(X) = \{0\} \cup \{(n, m) \mid m \in F(set(n))\}.$$

There is a homeomorphic embedding

$$Cont[\mathcal{P}(\mathbb{N}), \mathcal{P}(\mathbb{N})] \to \{\lambda X.F(X) \mid F \in \mathcal{P}(\mathbb{N})\},$$

where $Cont[\mathcal{P}(\mathbb{N}), \mathcal{P}(\mathbb{N})]$ denotes the space of continuous funtions on $\mathcal{P}(\mathbb{N})$ (w.r.t. positive topology). This gives a natural topology to the space of continuous functions, which is a retract of $\mathcal{P}(\mathbb{N})$.

It is useful to introduce a couple of continuous functions:

$$\mathbf{Pair}(X)(Y) = \{2n \mid n \in X\} \cup \{2m+1 \mid m \in Y\},$$
$$\mathbf{Fst}(Z) = \{n \mid 2n \in Z\}, \qquad \mathbf{Snd}(Z) = \{m \mid 2m+1 \in Z\},$$
$$\mathbf{Test}(Z)(X)(Y) = \{n \in X \mid 0 \in Z\} \cup \{m \in Y \mid \exists k.k+1 \in Z\}.$$

These definitions make the topological space $\mathcal{P}(\mathbb{N})$ homeomorphic to its cartesian square, $\mathcal{P}(\mathbb{N}) \cong \mathcal{P}(\mathbb{N}) \times \mathcal{P}(\mathbb{N})$ and show that subsets of $\mathcal{P}(\mathbb{N})$ can be regarded as binary relations on $\mathcal{P}(\mathbb{N})$.

Before concluding this paragraph we shall emphasize a few topological aspects connecting $\mathcal{P}(\mathbb{N})$ and $2^{\mathbb{N}}$ that will be of great use in what follows.

There exists a straightforward bijection $\chi : \mathcal{P}(\mathbb{N}) \to 2^{\mathbb{N}}$ that maps a subset of $\mathbb{N}$ to its characteristic function: $\chi(a)_i = 1$ if $i \in a$ and $\chi(a)_i = 0$ if $i \notin a$. The image of a basic open set of the Scott topology under χ is $\{\beta \mid \chi(a) \leq \beta\}$, where a is a finite set and $\leq$ is the componentwise extension of the order $0 \leq 1$. This space is the topological power of ω copies of the *Sierpiński space*, the two-element T_0 space with open sets $\emptyset$, $\{1\}$, and $\{0, 1\}$. Thus, we have two topologies on $2^{\mathbb{N}}$ but they both generate the same Borel sets $\mathcal{B}$, and hence, when equipped with the coin-flipping probability measure, the same Lebesgue completion. This follows from the fact that every basic Cantor open set is a finite Boolean combination of basic Scott open sets and vice versa [19].

3.2 Random Variables on Ω

Let $\Omega = (2^{\mathbb{N}}, \mathcal{B}, P)$ denote the standard probability space of infinite binary sequences, with $\mathcal{B}$ the Borel-algebra of the Cantor topology, and with the *coin-flipping* probability measure P, as defined in § 2. The following theorem relates Ω to the reals and allows us to use Ω as the "source of randomness" when we define random variables.

Theorem 3.3. *The measure spaces Ω and $[0, 1]$ with Lebesgue measure, restricted to Borel sets, are Borel isomorphic.*

We view $\mathcal{P}(\mathbb{N})$ as a measurable space with the Borel-algebra of the positive topology on $\mathcal{P}(\mathbb{N})$. This allows us to define *random variables* on $\Omega = (2^{\mathbb{N}}, \mathcal{B}, P)$ as the measurable functions

$$\xi : \Omega \to \mathcal{P}(\mathbb{N}).$$

Let $\mathcal{R}(\Omega)$ denote the set of these $\mathcal{P}(\mathbb{N})$-valued random variables on Ω. Given a random variable ξ, we obtain a measure on $\mathcal{P}(\mathbb{N})$ by

$P \circ \xi^{-1}$. A family of random variables is *independent* if the induced measures are independent.

On $\mathcal{R}(\Omega)$ we define a few functions. Firstly, we associate to any $X \in \mathcal{P}(\mathbb{N})$ a random variable $\hat{X}$ defined by $\hat{X}(\omega) = X$ for any $\omega \in \Omega$. Secondly, since application, λ-abstraction, **Pair**, **Fst** and **Snd** are all continuous functions on $\mathcal{P}(\mathbb{N})$, they can be canonically extended to $\mathcal{R}(\Omega)$. In [18] it is emphasized that $\mathcal{R}(\Omega)$ is a (non-extensional) model for the untyped λ-calculus. Moreover, an equation between two random variables $\xi, \eta \in \mathcal{R}(\Omega)$ can be interpreted as the *measurable event*

$$[\xi = \eta] = \{\omega \in \Omega \mid \forall n \in \xi(\omega).n \in \eta(\omega) \wedge \forall n \in \eta(\omega).n \in \xi(\omega)\},$$

which is a Borel set in $\mathcal{B}$, and this motivates the study of the algebra of events of Ω defined below.

3.3 The Algebra of Events

Given the probability space $\Omega = (2^{\mathbb{N}}, \mathcal{B}, P)$, we define the *algebra of events* (also known as the *measure algebra*) as the Boolean algebra

$$\mathcal{B}/_{Null} = (\mathcal{B}/_{Null}, \cup, \cap, \sim, \emptyset/_{Null}, \Omega/_{Null})$$

which is the quotient algebra of the σ-algebra $\mathcal{B}$ modulo the σ-ideal of Borel sets of P-measure zero. Observe that we do not get more expressive if we consider the P-Lebesgue completion of $\mathcal{B}/_{Null}$: because every Lebesgue measurable set differs from a Borel set by a null set, the measure algebra of Lebesgue measurable sets modulo P-null sets is isomorphic to $\mathcal{B}/_{Null}$.

We call the elements of $\mathcal{B}/_{Null}$ *events* and for $A \in \mathcal{B}$ we denote its equivalence class by $A/_{Null}$.

Theorem 3.4. *$(\mathcal{B}/_{Null}, \cup, \cap, \sim, \emptyset/_{Null}, \Omega/_{Null})$ is a σ-complete Boolean algebra in which any family of pairwise disjoint elements is countable (*i.e. *it satisfies the* countable chain condition*). Therefore it is a complete Boolean algebra.*

$\mathcal{B}/_{Null}$ plays a central role in the semantics of the stochastic λ-calculus.

4 Tossing processes

A key ingredient in any probabilistic programming language is the source of randomness. As in [18], this is taken to be a random variable which uses an infinite sequence of independent fair coin tosses to resolve the random choice. The semantics should not depend on the vagaries of a particular sequence; accordingly, we aim to prove a property that shows that the semantics should be independent, in a suitable sense, of the coin tosses that occur. This is where the notion of *monolithic function* becomes important. We call the special random variables that we use *tossing processes*. This section is devoted to the properties of tossing processes.

4.1 Independent coin sequences

We need to move between sequences of coin tosses and subsets of $\mathbb{N}$ by using some appropriate coding and decoding functions. Concretely, we can a Borel-measurable map $\mathbf{pack} : 2^{\mathbb{N}} \to \mathcal{P}(\mathbb{N})$ which encodes a sequence as a set in a way that can be easily inverted and its inverse, called **unpack**, is Borel-measurable as well.

Moreover, these can be defined so that they properly relate the well-known operations on sequences $\mathbf{head} \colon 2^{\mathbb{N}} \to 2$ and $\mathbf{tail} \colon 2^{\mathbb{N}} \to 2^{\mathbb{N}}$, given by $\mathbf{head}(a) = a_0$ and $\mathbf{tail}(a)(n) = a_{n+1}$, to the functions **Fst** and **Snd** defined on sets in Section 3 (*i.e.*, on constant random variables) as stated in the following lemma[7].

Lemma 4.1. *Let $a \in 2^{\mathbb{N}}$ and $\alpha \in \{0, 1\}$.*

1. *$\{\alpha\} = \mathbf{Fst}(\mathbf{pack}(a))$ iff $\mathbf{head}(a) = \alpha$.*
2. *$\mathbf{Snd}(\mathbf{pack}(a)) = \mathbf{pack}(\mathbf{tail}(a))$.*

Let $\Omega = (2^{\mathbb{N}}, \mathcal{B}, P)$ be the probability space defined in § 3.2.

Definition 4.2. A *coin flip* is a random variable that has the form $F \colon \Omega \to \{\{0\}, \{1\}\}$. A coin flip is *fair* whenever $P(F^{-1}(\{0\})) = 1/2$.

An *independent sequence of coin tosses* (ICS) is a random variable $\mathcal{T} \colon \Omega \to \mathcal{P}(\mathbb{N})$ such that $\mathbf{Fst}(\mathcal{T})$ is a fair coin flip and $\mathbf{Snd}(\mathcal{T})$ is another ICS—with the successive flips all mutually independent.

Note that an ICS is a $\{\{0\}, \{1\}\}$-valued map, i.e., specialized to take values in the image of **pack**. Every ICS $\mathcal{T} \colon \Omega \to \mathcal{P}(\mathbb{N})$ is of the form $\mathcal{T} = \mathbf{pack} \circ T$ for some (Borel) measurable map $T \colon \Omega \to 2^{\mathbb{N}}$. By Lemma 4.1, $\mathbf{Fst} \circ \mathbf{pack} \circ T = \mathbf{head} \circ T$, so the condition of $\mathbf{Fst}(\mathcal{T})$ being a fair coin is that for $i \in \{0, 1\}$,

$$P((\mathbf{head} \circ T)^{-1}(i)) = \frac{1}{2},$$

That $\mathbf{Snd}(\mathcal{T})$ is an ICS implies that for all $n \in \mathbb{N}$ and $i \in \{0, 1\}$,

$$P((\mathrm{ev}_n \circ T)^{-1}(i)) = \frac{1}{2},$$

where $\mathrm{ev}_n \colon 2^{\mathbb{N}} \to \{0, 1\}$ for $n \in \mathbb{N}$, is defined by $\mathrm{ev}_n(a) = a_n$.

The condition of independence implies that for any finite increasing sequence $(m_i)_{i=1}^n$ in $\mathbb{N}$, and finite sequence $(b_i)_{i=1}^n$ in $\{0, 1\}$,

$$P\left(\bigcap_{i=1}^{n} (\mathrm{ev}(m_i) \circ T)^{-1}(b_i)\right) = 2^{-n}.$$

This means that the image measure $T_*(P) = P \circ T^{-1}$ agrees with the standard coin-flipping measure on $2^{\mathbb{N}}$ on basic clopens of Cantor topology. As basic clopens form a π-system (they are closed under finite intersections) and generate the Borel sets of $2^{\mathbb{N}}$, $T_*(P)$ and the coin-flipping measure agree on all Borel sets [22, Lemma 1.6], so we have that T is measure-preserving from Ω to $2^{\mathbb{N}}$.

Conversely, for any measure-preserving map $T \colon \Omega \to 2^{\mathbb{N}}$, $\mathbf{pack} \circ T$ will be an independent sequence of coin tosses.

These provide the following characterization for ICS.

Theorem 4.3 (Characterization)**.** *Independent coin sequences are exactly the maps of the form* $\mathbf{pack} \circ T$*, for some measure-preserving map $T \colon \Omega \to \Omega$.*

Now we focus on another important function for our discussion,

$$\mathbf{evens} : 2^{\mathbb{N}} \to 2^{\mathbb{N}},$$

defined by $\mathbf{evens}(a)(n) = a_{2n}$.

[7] A detailed construction of a pair of such coding/decoding functions can be found in the appendix.

Theorem 4.4. *A measure-preserving map $T : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ is monolithic iff $T \cong$ **evens**.*

Proof. Because $\mathbb{N} \cong \mathbb{N} + \mathbb{N}$ by mapping the odd numbers to the first part and the even numbers to the second part, we have $(2^{\mathbb{N}} \times 2^{\mathbb{N}}, \pi_2) \cong (2^{\mathbb{N}}, \mathbf{evens})$. As any atomless standard probability space is isomorphic to $2^{\mathbb{N}}$, applying Theorem 2.4 with X and Y specialized to $2^{\mathbb{N}}$, we get the desired equivalence. □

4.2 Tossing Processes

We are ready to define the concept of tossing process; the crucial point is to insist on a monolithic function.

Definition 4.5 (Tossing Process)**.** A *tossing process* is an independent sequence of coin tosses $\mathcal{T} = \mathbf{pack} \circ T$, where $T : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ is a monolithic measure-preserving map.

We denote by **Toss** the set of tossing processes. The next theorem states that all tossing processes are the same up to a measure-preserving automorphism of the measure algebra. Let $Aut(\Omega)$ denote the set of measure-preserving automorphisms on Ω.

Theorem 4.6 (Representation Theorem)**.** *For any two tossing processes $\mathcal{S}$ and $\mathcal{T}$ there exists a measure-preserving automorphism $\alpha : \Omega \to \Omega \in Aut(\Omega)$ such that $\mathcal{T} = \mathcal{S} \circ \alpha$ except for on a subset of Ω of measure 0.*

4.3 Tossing Process Operators

Before concluding this section, we show a few useful operators on $\mathcal{R}(\Omega)$ that are closed on tossing processes.

Let

$$(.)^e, (.)^o, \mathbf{Swap} : \mathbf{Toss} \to \mathcal{R}(\Omega),$$

defined for an arbitrary tossing process $\mathcal{T} = \mathbf{pack} \circ T$ as follows.

$$\mathcal{T}^e = \mathbf{pack} \circ \mathbf{evens} \circ T$$
$$\mathcal{T}^o = \mathbf{pack} \circ \mathbf{odds} \circ T$$
$$\mathbf{Swap}(\mathcal{T}) = \mathbf{pack} \circ \mathit{swap} \circ T,$$

where $\mathbf{odds}, \mathit{swap} \colon \Omega \to \Omega$ are defined by $\mathbf{odds}(a)(n) = a_{2n+1}$ and $\mathit{swap}(a)(2n) = a_{2n+1}$, $\mathit{swap}(a)(2n+1) = a_{2n}$.

Lemma 4.7. *If $\mathcal{T}$ is a tossing process, then* $\mathbf{Snd}(\mathcal{T})$, $\mathcal{T}^e$, T^o *and* $\mathbf{Swap}(\mathcal{T})$ *are tossing processes as well.*

5 Stochastic λ-Calculus

In this section we introduce the stochastic λ-calculus. The syntax of the stochastic λ-calculus over a set $V \ni x$ of variables extends the syntax of the classical untyped λ-calculus with a (binary) probabilistic choice operator $\oplus$ on λ-terms, and a fixpoint combinator $\mu x.M$:

$$M := x \mid \lambda x.M \mid MM \mid \mu x.M \mid M \oplus M.$$

Notation: In what follows we will call terms without any occurrence of $\oplus$ *stable terms*; however, a stable term may contain the fixpoint operator. And, as usual, we use $M\{N/x\}$ to denote the substitution of the variable x by the term N in M.

Unlike in the classical λ-calculus, in the stochastic λ-calculus the fixpoint combinator cannot be defined from the other operators, as we will demonstrate later. For the development of the fixpoint operator, it is useful to define the *unfolding* of a recursive term.

Given $\mu x.M$, its *unfolding* is the sequence of terms $M^0, M^1, \ldots$ defined inductively as follows.

$$M^0 = (\lambda x\,.\,xx)(\lambda x\,.\,xx), \quad \text{and for arbitrary } n,\ M^{n+1} = (\lambda x\,.\,M)M^n.$$

6 Probabilistic Continuation Semantics

The continuation semantics for the stochastic λ-calculus interprets a λ-term relative to an *environment* giving values to the free variables, a *continuation* giving a subsequent computation, and a *tossing process* $\mathcal{T}$ used to resolve probabilistic choices.

We know that the set $\mathcal{R}(\Omega)$ of random variables with the pointwise order forms a domain such that $[\mathcal{R}(\Omega) \to \mathcal{R}(\Omega)]$ is a continuous retract of $\mathcal{R}(\Omega)$, where $[\mathcal{R}(\Omega) \to \mathcal{R}(\Omega)]$ is the space of Scott-continuous functions. We write this explicitly by introducing the functions

$$\phi : \mathcal{R}(\Omega) \to [\mathcal{R}(\Omega) \to \mathcal{R}(\Omega)] \text{ and } \psi : [\mathcal{R}(\Omega) \to \mathcal{R}(\Omega)] \to \mathcal{R}(\Omega).$$

It is useful to also define the direct and the continuation-passing semantics for stable λ-terms. We use x to range over variables, E to range over arbitrary environments, C to range over arbitrary continuations, and $\mathcal{T}$ to range over arbitrary tossing processes. As before, the n-th unfolding of the term $\mu x.M$ is denoted by M^n.

Let $\langle\cdot\rangle$ denote the direct semantics and $\langle\!\langle\cdot\rangle\!\rangle$ the continuation-passing semantics for stable λ-terms. In addition, for stochastic terms, we denote by $(\!|\cdot|\!)$ the continuation-passing semantics augmented with a tossing process.

$$\langle\cdot\rangle : \mathbf{Term} \to \mathbf{Env} \to \mathcal{R}(\Omega)$$
$$\langle\!\langle\cdot\rangle\!\rangle : \mathbf{Term} \to \mathbf{Env} \to \mathbf{Cont} \to \mathcal{R}(\Omega)$$
$$(\!|\cdot|\!) : \mathbf{Term} \to \mathbf{Env} \to \mathbf{Cont} \to \mathbf{Toss} \to \mathcal{R}(\Omega).$$

where

$$\mathbf{Env} = \mathbf{Var} \to \mathcal{R}(\Omega) \qquad \mathbf{Cont} = [\mathcal{R}(\Omega) \to \mathcal{R}(\Omega)]$$

The **direct semantics** is

$$\langle x\rangle E = E(x)$$
$$\langle MN\rangle E = \phi(\langle M\rangle E)(\langle N\rangle E)$$
$$\langle \lambda x\,.\,M\rangle E = \psi(\lambda v\,.\,\langle M\rangle(E\{v/x\}))$$
$$\langle \mu x.M\rangle E = \sup_n \langle M^n\rangle E.$$

With **continuations**, define

$$\langle\!\langle x\rangle\!\rangle EC = C(E(x))$$
$$\langle\!\langle MN\rangle\!\rangle EC = \langle\!\langle M\rangle\!\rangle E(\lambda a\,.\,\langle\!\langle N\rangle\!\rangle E(\lambda b\,.\,C(\phi(a)b)))$$
$$\langle\!\langle \lambda x\,.\,M\rangle\!\rangle EC = C(\psi(\lambda v\,.\,\langle\!\langle M\rangle\!\rangle(E\{v/x\})(\lambda u\,.\,u)))$$
$$\langle\!\langle \mu x.M\rangle\!\rangle EC = \sup_n \langle\!\langle M^n\rangle\!\rangle EC.$$

The **probabilistic continuation** is defined as follows.

$$(\!|x|\!)EC\mathcal{T} = C(E(x))$$
$$(\!|MN|\!)EC\mathcal{T} = (\!|M|\!)E(\lambda a\,.\,(\!|N|\!)E(\lambda b\,.\,C(\phi(a)b))\mathcal{T}^e)\mathcal{T}^o$$
$$(\!|\lambda x\,.\,M|\!)EC\mathcal{T} = C(\psi(\lambda v\,.\,(\!|M|\!)(E\{v/x\})(\lambda u\,.\,u)\mathcal{T}))$$
$$(\!|M \oplus N|\!)EC\mathcal{T} = \mathbf{Test}(\mathbf{Fst}(\mathcal{T}))((\!|M|\!)EC(\mathbf{Snd}(\mathcal{T})))((\!|N|\!)EC(\mathbf{Snd}(\mathcal{T})))$$
$$(\!|\mu x.M|\!)EC\mathcal{T} = \sup_n (\!|M^n|\!)EC\mathcal{T}$$

The relation between the three semantics for stable terms is stated in the following proposition.

Proposition 6.1. *If M is a stable term, then for an arbitrary environment E, an arbitrary continuation process C, and an arbitrary tossing process $\mathcal{T}$,*

$$C(\langle M\rangle E) = \langle\!\langle M\rangle\!\rangle EC = (\!|M|\!)EC\mathcal{T}\,.$$

A corollary of this lemma is that if a closed program has a value, then its value is the same for all tossing processes.

Corollary 6.2 (Absoluteness I)**.** *If M is a stable term, then for an arbitrary environment E, continuation process C, and tossing processes $\mathcal{T}, \mathcal{T}'$,*

$$(\!|M|\!)EC\mathcal{T} = (\!|M|\!)EC\mathcal{T}'\,.$$

We conclude this section with two useful lemmas.

Lemma 6.3. *For any stochastic λ-terms M, N, any arbitrary environment E, any arbitrary continuation C and any arbitrary tossing process $\mathcal{T}$, the following statements hold, where M^n denote the n-unfolding of $\mu x.M$.*

1. $(\!|\lambda y.(\mu x.M)|\!)EC\mathcal{T} = \sup_n (\!|\lambda y.M^n|\!)EC\mathcal{T}$;
2. $(\!|N(\mu x.M)|\!)EC\mathcal{T} = \sup_n (\!|NM^n|\!)EC\mathcal{T}$;
3. $(\!|(\lambda x.M)N|\!)EC\mathcal{T} = \sup_n (\!|M^n N|\!)EC\mathcal{T}$;

Lemma 6.4. *For any stochastic λ-terms M, N, any arbitrary environment E, any arbitrary continuation C and any arbitrary tossing process $\mathcal{T}$,*

$$(\!|(\lambda x.M)N|\!)EC\mathcal{T} = (\!|M|\!)E\{(\!|N|\!)E(\lambda w.w)\mathcal{T}^e/x\}C\mathcal{T}^o.$$

We conclude this section by presenting a direct consequence of Theorem 4.6.

Theorem 6.5 (Absoluteness II)**.** *Given a term M, for an arbitrary environment E, an arbitrary continuation process C, and arbitrary tossing processes $\mathcal{T}, \mathcal{T}'$, there exists a measure-preserving automorphism $\alpha : \Omega \to \Omega \in Aut(\Omega)$ such that*

$$\{\omega \in \Omega \mid (\!|M|\!)EC\mathcal{T}(\omega) = (\!|M|\!)EC(\mathcal{T}' \circ \alpha)(\omega)\}/_{Null} = \Omega/_{Null}\,.$$

7 A Boolean-Valued Model

The Boolean-valued model gives a novel interpretation of equality. Equalities of closed terms, when interpreted over $\mathcal{R}(\Omega)$, are associated with events in $\mathcal{B}/_{Null}$ up to a measure-preserving automorphism of Ω. Since we are working with closed terms we evaluate terms in the empty environment and with the the identity continuation.

Let $\emptyset$ denote the empty environment and $id = \lambda x.x$ denote the identity continuation.

Definition 7.1. For arbitrary closed terms M, N, and tossing process $\mathcal{T}$, let

$$[\![M = N]\!]_{\mathcal{T}} = \left\{\omega \in \Omega \;\middle|\; (\!|M|\!)\emptyset id\mathcal{T}(\omega) = (\!|N|\!)\emptyset id\mathcal{T}(\omega)\right\}/_{Null}\,.$$

Note that $[\![M = N]\!]_{\mathcal{T}} \in \mathcal{B}/_{Null}$ and that this value depends directly on the tossing process $\mathcal{T}$. However, since the tossing processes are all equal up to a measure-preserving automorphism, as proven in Theorem 4.6, $[\![M = N]\!]_{\mathcal{T}}$ is unique up to an automorphism of $\mathcal{B}/_{Null}$.

In what follows, for arbitrary $A, A' \in \mathcal{B}/_{Null}$, we write

$$A \approx A'$$

if there exists a measure-preserving automorphism σ of $\mathcal{B}/_{Null}$ such that $\sigma(A) = A'$.

Theorem 7.2 (Absoluteness III)**.** *For arbitrary closed terms M, N, and arbitrary tossing processes $\mathcal{T}$ and $\mathcal{T}' = \mathcal{T} \circ \alpha$, where $\alpha \in Aut(\Omega)$,*

$$\alpha^{-1}([\![M = N]\!]_{\mathcal{T}}) = [\![M = N]\!]_{\mathcal{T}'},$$

where α^{-1} is the set-theoretical inverse of α, hence an automorphism of $\mathcal{B}/_{Null}$. Consequently,

$$[\![M = N]\!]_{\mathcal{T}} \approx [\![M = N]\!]_{\mathcal{T}'}.$$

Proof. $[\![M = N]\!]_{\mathcal{T}'} = \left\{\omega \in \Omega \;\middle|\; (\!|M|\!)\emptyset id\mathcal{T}'(\omega) = (\!|N|\!)\emptyset id\mathcal{T}'(\omega)\right\}/_{Null}$

$= \left\{\omega \in \Omega \;\middle|\; (\!|M|\!)\emptyset id(\mathcal{T} \circ \alpha)(\omega) = (\!|N|\!)\emptyset id(\mathcal{T} \circ \alpha)(\omega)\right\}/_{Null}.$

Since $\emptyset$ and id are constant, this set is further equal to

$$\begin{aligned}&\left\{\omega \in \Omega \;\middle|\; (\!|M|\!)\emptyset id\mathcal{T}(\alpha(\omega)) = (\!|N|\!)\emptyset id\mathcal{T}(\alpha(\omega))\right\}/_{Null}\\ &= \alpha^{-1}(\left\{\omega \in \Omega \;\middle|\; (\!|M|\!)\emptyset id\mathcal{T}(\omega) = (\!|N|\!)\emptyset id\mathcal{T}(\omega)\right\}/_{Null})\\ &= \alpha^{-1}([\![M = N]\!]_{\mathcal{T}})\,.\end{aligned}$$

□

This last theorem suggests that in what follows we can use any tossing process to evaluate the equality between closed programs, since the result is in any case unique up to an automorphism of the measure algebra.

8 Sound Equations

In this section we establish a series of sound equations that provide basic reasoning principles for our stochastic λ-calculus. These equations are by no means complete, but they do describe several basic facts about $\mathcal{R}(\Omega)$. We also show how the usual equations for α-reduction and β-reduction are generalized. Note that, for stable closed terms, we have the soundness of α-reduction and β-reduction for the model $\mathcal{R}(\Omega)$ of ordinary λ-calculus from Proposition 6.1.

In what follows, all the terms in expressions of the form $[\![M = N]\!]_{\mathcal{T}}$ are implicitly assumed to be closed terms.

The first result shows how one can substitute terms in equations with terms that are equal *almost everywhere*.

Theorem 8.1 (Substitution)**.** *If M and N are closed terms such that $[\![M = N]\!]_{\mathcal{T}} = \Omega/_{Null}$, then for any closed stable term K,*

$$[\![M = K]\!]_{\mathcal{T}} = [\![N = K]\!]_{\mathcal{T}}.$$

Proof. We have that

$$[\![M = N]\!]_{\mathcal{T}} = \{\omega \in \Omega \mid (\!|M|\!)\emptyset id\mathcal{T}(\omega) = (\!|N|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null}.$$

This means that $[\![M = N]\!]_{\mathcal{T}} = \Omega/_{Null}$ implies that $(\!|M|\!)\emptyset id\mathcal{T}$ and $(\!|N|\!)\emptyset id\mathcal{T}$ are equal almost everywhere. But then,

$$[\![M = K]\!]_{\mathcal{T}} = \{\omega \in \Omega \mid (\!|M|\!)\emptyset id\mathcal{T}(\omega) = (\!|K|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null}$$
$$= \{\omega \in \Omega \mid (\!|N|\!)\emptyset id\mathcal{T}(\omega) = (\!|K|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null} = [\![N = K]\!]_{\mathcal{T}}. \quad \square$$

There are many interesting properties that one can prove in this setting. We begin by observing that both α-reduction and β-reduction for stable terms (*i.e.* terms without any occurence of $\oplus$) hold, as direct consequences of the fact that $\mathcal{R}(\Omega)$ is a model of the usual untyped λ-calculus [18]. However, α-reduction also holds for our stochastic λ-calculus; this follows from the probabilistic continuation semantics of λ-terms introduced in Section 6.

Theorem 8.2 (α-reduction)**.** *If M is a term without free occurrences of y, then*

$$[\![\lambda x.M = \lambda y.M\{y/x\}]\!]_{\mathcal{T}} = \Omega/_{Null}.$$

Now we state β-reduction only for stable terms. Later in this section we will show that an unrestricted version of β-reduction is not possible, but that we have, however, some extensions that involve terms that might be not stable.

Theorem 8.3 (β-reduction)**.** *Let N and M be stable terms. Then,*

$$[\![(\lambda x.M)(N) = M\{N/x\}]\!]_{\mathcal{T}} = \Omega/_{Null}.$$

Next we prove a series of results regarding the properties of the probabilistic choice operator. We start with the following lemma.

Lemma 8.4. *Let M_1, M_2, N_1, N_2 be closed terms and $\mathcal{T}$ an arbitrary tossing process. If for any stable closed term K and each $i \in \{1, 2\}$ we have that $[\![M_i = K]\!]_{\mathbf{Snd}(\mathcal{T})} \approx [\![N_i = K]\!]_{\mathbf{Snd}(\mathcal{T})}$, then for any stable closed term K,*

$$[\![M_1 \oplus M_2 = K]\!]_{\mathcal{T}} \approx [\![N_1 \oplus N_2 = K]\!]_{\mathcal{T}}.$$

Proof. The key observation is that there exists a $j \in \{1, 2\}$ such that

$$(\!|M_1 \oplus M_2|\!)\emptyset id\mathcal{T} = (\!|M_j|\!)\emptyset id\mathbf{Snd}(\mathcal{T})$$

and at the same time

$$(\!|N_1 \oplus N_2|\!)\emptyset id\mathcal{T} = (\!|N_j|\!)\emptyset id\mathbf{Snd}(\mathcal{T}).$$

Also, since K is stable, $(\!|K|\!)\emptyset id\mathcal{T} = (\!|K|\!)\emptyset id\mathbf{Snd}(\mathcal{T})$. Hence,

$$[\![M_1 \oplus M_2 = K]\!]_{\mathcal{T}}$$
$$= \{\omega \in \Omega \mid (\!|M_1 \oplus M_2|\!)\emptyset id\mathcal{T}(\omega) = (\!|K|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null}$$
$$= \{\omega \in \Omega \mid (\!|M_j|\!)\emptyset id\mathbf{Snd}(\mathcal{T})(\omega) = (\!|K|\!)\emptyset id\mathbf{Snd}(\mathcal{T})(\omega)\}/_{Null}$$
$$= [\![M_j = K]\!]_{\mathbf{Snd}(\mathcal{T})}.$$

Similarly, $[\![N_1 \oplus N_2 = K]\!]_{\mathcal{T}} = [\![N_j = K]\!]_{\mathbf{Snd}(\mathcal{T})}$. Now the hypothesis guarantees that $[\![M_j = K]\!]_{\mathbf{Snd}(\mathcal{T})} \approx [\![N_j = K]\!]_{\mathbf{Snd}(\mathcal{T})}$. $\square$

A first fundamental property of probabilistic choice is a kind of commutativity, stated in the following axiom. Note that one cannot assert commutativity naively; one has to talk in terms of a stable term obtained by resolving all the choices.

Theorem 8.5 ($\oplus$-commutativity)**.** *If K is a stable term, then we have*

$$[\![M \oplus N = K]\!]_{\mathcal{T}} \approx [\![N \oplus M = K]\!]_{\mathcal{T}}.$$

Proof. The map neg : $2^{\mathbb{N}} \to 2^{\mathbb{N}}$, where $\mathrm{neg}(a)(0) = 1 - a(0)$ and $\mathrm{neg}(a)(i) = a(i)$ for $i > 0$ is measurable and measure-preserving. The tossing process $\mathcal{T}$ is of the form $\mathbf{pack} \circ T$ for some measure-preserving monolithic map $T : \Omega \to 2^{\mathbb{N}}$. By Lemma 2.5, $\mathcal{T}' = \mathbf{pack} \circ \mathrm{neg} \circ T$ is also a tossing process, and so by Theorem 4.6 there exists a measure-preserving automorphism $\alpha : \Omega \to \Omega$ such that $\mathcal{T}' = \alpha \circ \mathcal{T}$. We start by observing that

$$\begin{aligned}
&(\!|M \oplus N|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T} \\
&= \mathbf{Test}(\mathbf{Fst}(\mathcal{T}))((\!|M|\!)\,\emptyset\,\mathrm{id}\,(\mathbf{Snd}(\mathcal{T})))((\!|N|\!)\,\emptyset\,\mathrm{id}\,(\mathbf{Snd}(\mathcal{T}))) \\
&= \mathbf{Test}(\mathbf{Fst}(\mathcal{T}'))((\!|N|\!)\,\emptyset\,\mathrm{id}\,(\mathbf{Snd}(\mathcal{T})))((\!|M|\!)\,\emptyset\,\mathrm{id}\,(\mathbf{Snd}(\mathcal{T}))) \\
&= (\!|N \oplus M|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}' \\
&= (\!|N \oplus M|\!)\,\emptyset\,\mathrm{id}\,(\mathcal{T} \circ \alpha),
\end{aligned}$$

using the fact that $\mathbf{Snd}(\mathcal{T}') = \mathbf{Snd}(\mathcal{T})$, since neg only affects the first part. So

$$\begin{aligned}
&[\![M \oplus N = K]\!]_{\mathcal{T}} \\
&= \{\omega \in \Omega \mid (\!|M \oplus N|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}(\omega) = (\!|K|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}(\omega)\}/_{Null} \\
&= \{\omega \in \Omega \mid (\!|N \oplus M|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}(\alpha(\omega)) = (\!|K|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}(\omega)\}/_{Null} \\
&= \{\omega \in \Omega \mid (\!|N \oplus M|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}(\alpha(\omega)) = (\!|K|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}(\alpha(\omega))\}/_{Null} \\
&= \alpha^{-1}(\{\omega \in \Omega \mid (\!|N \oplus M|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}(\omega) = (\!|K|\!)\,\emptyset\,\mathrm{id}\,\mathcal{T}(\omega)\})/_{Null} \\
&= \alpha^{-1}([\![N \oplus M = K]\!]),
\end{aligned}$$

where the third equality holds because $(\!|K|\!)$ does not depend on ω since it is stable. $\square$

Next we state that $\oplus$ is idempotent in the same sense as in the previous theorem.

Theorem 8.6 ($\oplus$-idempotence)**.** *If K is a stable term, then we have*

$$[\![M \oplus M = K]\!]_{\mathcal{T}} \approx [\![M = K]\!]_{\mathcal{T}}.$$

The next two theorems state that $\oplus$ is distributive to the left and to the right with respect to application.

Theorem 8.7 (Left-distributivity of $\oplus$ w.r.t. application)**.** *If K is a stable term, then we have*

$$[\![(M_1 \oplus M_2)(N) = K]\!]_{\mathcal{T}} \approx [\![M_1 N \oplus M_2 N = K]\!]_{\mathcal{T}}.$$

Proof. The semantics of application gives us

$$(\!|(M_1 \oplus M_2)(N)|\!)\emptyset id\mathcal{T} = (\!|M_1 \oplus M_2|\!)\emptyset(\lambda x.(\!|N|\!)\emptyset(\lambda y.(xy\ id))\mathcal{T}^e)\mathcal{T}^o.$$

Assume that $\mathbf{Test}(\mathbf{Fst}(\mathcal{T}))$, applied to a particular $\omega \in \Omega$ chooses M_i for some $i \in \{1, 2\}$, in the continuation semantics of $M_1 \oplus M_2$.

Let $\mathcal{T}' = \mathbf{Snd}(\mathbf{Swap}(\mathcal{T}))$. The previous term is further equal to

$$\begin{aligned}
(\!|M_i|\!)\emptyset(\lambda x.(\!|N|\!)\emptyset(\lambda y.(xy\ id))\mathcal{T}^e)(\mathbf{Snd}(\mathcal{T}^o)) &= (\!|M_i N|\!)\emptyset id(\mathbf{Snd}(\mathcal{T}')) \\
&= (\!|M_1 N \oplus M_2 N|\!)\emptyset id\mathcal{T}'.
\end{aligned}$$

In the last line we have used the fact that if $\mathbf{Test}(\mathbf{Fst}(\mathcal{T}))$, when applied to a particular $\omega \in \Omega$ in the semantics of $M_1 \oplus M_2$, chooses M_i, then the same test applied to the same ω in the semantics of $M_1 N \oplus M_2 N$ will choose $M_i N$ for the same $i \in \{1, 2\}$.

Since **Snd** and **Swap** both preserve tossing processes (Lemma 4.7), there exists an automorphism $\alpha \in Aut(\Omega)$ such that $\mathcal{T}' = \mathcal{T} \circ \alpha$. Hence, we have

$$[\![(M_1 \oplus M_2)(N) = K]\!]_{\mathcal{T}}$$
$$= \{\omega \in \Omega \mid (\!|(M_1 \oplus M_2)N|\!)\emptyset id\mathcal{T}(\omega) = (\!|K|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null}$$

$$= \{\omega \in \Omega \mid (\!|M_1 N \oplus M_2 N|\!)\emptyset id\mathcal{T}'(\omega) = (\!|K|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null}.$$

Since K is stable, $(\!|K|\!)\emptyset id\mathcal{T}(\omega) = (\!|K|\!)\emptyset id(\mathcal{T}\circ\alpha)(\omega)$ and the previous set is equal to

$$= \{\omega \in \Omega \mid (\!|M_1 N \oplus M_2 N|\!)\emptyset id\mathcal{T}'(\omega) = (\!|K|\!)\emptyset id\mathcal{T}'(\omega)\}/_{Null}$$
$$= \alpha^{-1}(\{\omega \in \Omega \mid (\!|M_1 N \oplus M_2 N|\!)\emptyset id\mathcal{T}(\omega) = (\!|K|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null})$$
$$= \alpha^{-1}([\![M_1 N \oplus M_2 N = K]\!]_{\mathcal{T}}).$$ □

Theorem 8.8 (Right-distributivity of $\oplus$ w.r.t. application)**.** *If K is a stable term, then we have*

$$[\![N(M_1 \oplus M_2) = K]\!]_{\mathcal{T}} \approx [\![NM_1 \oplus NM_2 = K]\!]_{\mathcal{T}}.$$

Proof. The proof is similar to the one for left-distributivity, except that instead of $\mathcal{T}'$ we use $\mathcal{T}'' = \mathbf{Swap}(\mathbf{Snd}(\mathcal{T}))$. □

Using Lemma 8.4, we can prove the entropic equality equation.

Theorem 8.9 (Entropic equality)**.** *If K is a stable term, then we have*

$$[\![(M_1 \oplus M_2) \oplus (N_1 \oplus N_2) = K]\!]_{\mathcal{T}} \approx [\![(M_1 \oplus N_1) \oplus (M_2 \oplus N_2) = K]\!]_{\mathcal{T}}.$$

By exploiting the idempotence of $\oplus$ and the entropic equality, one can derive $\oplus$-distributivity.

Theorem 8.10 ($\oplus$-distributivity)**.** *If K is a stable term, then we have*

1. $[\![N \oplus (M_1 \oplus M_2) = K]\!]_{\mathcal{T}} \approx [\![(N \oplus M_1) \oplus (N \oplus M_2) = K]\!]_{\mathcal{T}}$;
2. $[\![(M_1 \oplus M_2) \oplus N = K]\!]_{\mathcal{T}} \approx [\![(M_1 \oplus N) \oplus (M_2 \oplus N) = K]\!]_{\mathcal{T}}$.

There are also some equalities between terms that are much stronger; these are equalities that hold almost everywhere which means that they are interpreted as the top element of the Boolean algebra. One such equation is λ-distributivity.

Theorem 8.11 (λ-distributivity w.r.t. $\oplus$)**.**

$$[\![\lambda x.(M_1 \oplus M_2) = \lambda x.M_1 \oplus \lambda x.M_2]\!]_{\mathcal{T}} = \Omega/_{Null}.$$

Proof. From the semantics of λ-terms we get

$$(\!|\lambda x.(M_1 \oplus M_2)|\!)\emptyset id\mathcal{T} = id(\psi(\lambda v\,.\,(\!|M_1 \oplus M_2|\!)(\emptyset\{v/x\})(\lambda u\,.\,u)\mathcal{T}))\,.$$

Assume that $\mathbf{Test}(\mathbf{Fst}(\mathcal{T}))$, applied to a particular $\omega \in \Omega$ chooses M_i for some $i \in \{1,2\}$, when applied in the continuation semantics of $M_1 \oplus M_2$. Then, the previous term is equal to

$$id(\psi(\lambda v\,.\,(\!|M_i|\!)(\emptyset\{v/x\})(\lambda u\,.\,u)(\mathbf{Snd}(\mathcal{T})))) = (\!|\lambda x.M_i|\!)\emptyset id(\mathbf{Snd}(\mathcal{T}))$$
$$= (\!|\lambda x.M_1 \oplus \lambda x.M_2|\!)\emptyset id\mathcal{T}\,.$$

In the last line we have used the fact that if $\mathbf{Test}(\mathbf{Fst}(\mathcal{T}))$, when applied to a particular $\omega \in \Omega$ in the semantics of $M_1 \oplus M_2$, chooses M_i, then the same test applied to the same ω in the semantics of $\lambda x.M_1 \oplus \lambda x.M_2$ will choose $\lambda x.M_i$ for the same $i \in \{1,2\}$. □

Theorem 8.12 (Order of applications)**.** *If N_1, N_2 are two stable closed terms, then*

$$[\![((\lambda x.\lambda y.M)N_1)N_2 = ((\lambda y.\lambda x.M)N_2)N_1]\!]_{\mathcal{T}} = \Omega/_{Null}.$$

At this point we are ready to prove the soundness of some equations involving the fixpoint operators.

Theorem 8.13 (Recursive application)**.**

$$[\![\mu x.M = (\lambda x.M)(\mu x.M)]\!]_{\mathcal{T}} = \Omega/_{Null}.$$

Proof. We know that, for each $n \geq 0$, $M^{n+1} = (\lambda x.M)M^n$. Hence, for arbitrary E, C and $\mathcal{T}$,

$$(\!|M^{n+1}|\!)EC\mathcal{T} = (\!|(\lambda x.M)M^n|\!)EC\mathcal{T}.$$

Lemma 6.4 applied to this equality gives us further that

$$(\!|M^{n+1}|\!)EC\mathcal{T} = (\!|M|\!)E\{(\!|M^n|\!)E(\lambda w.w)\mathcal{T}^e/x\}C\mathcal{T}^o.$$

Hence, $\sup_n (\!|M^{n+1}|\!)EC\mathcal{T} = \sup_n (\!|M|\!)E\{(\!|M^n|\!)E(\lambda w.w)\mathcal{T}^e/x\}C\mathcal{T}^o$ and using Scott continuity we get

$$\sup_n (\!|M^{n+1}|\!)EC\mathcal{T} = (\!|M|\!)E\{\sup_n (\!|M^n|\!)E(\lambda w.w)\mathcal{T}^e/x\}C\mathcal{T}^o.$$

Since M is continuous, hence monotonic, the above is equivalent to

$$\sup_n (\!|M^n|\!)EC\mathcal{T} = (\!|M|\!)E\{\sup_n (\!|M^n|\!)E(\lambda w.w)\mathcal{T}^e/x\}C\mathcal{T}^o,$$

or equivalently, $(\!|\mu x.M|\!)EC\mathcal{T} = (\!|M|\!)E\{(\!|\mu x.M|\!)E(\lambda w.w)\mathcal{T}^e/x\}C\mathcal{T}^o$ and again applying Lemma 6.4, $(\!|\mu x.M|\!)EC\mathcal{T} = (\!|(\lambda x.M)(\mu x.M)|\!)EC\mathcal{T}$. In particular, we also have

$$(\!|\mu x.M|\!)\emptyset id\mathcal{T} = (\!|(\lambda x.M)(\mu x.M)|\!)\emptyset id\mathcal{T}.$$

□

Theorem 8.14 (Recursive choice)**.**

$$[\![\mu x.(x \oplus M) \equiv \mu x.M]\!]_{\mathcal{T}} = \Omega/_{Null}.$$

Proof. Consider arbitrary E, C and $\mathcal{T}$. For each $\omega \in \Omega$, the sequence $(\!|M^i|\!)EC\mathcal{T}(\omega)$ is increasing and its limit is $(\!|\mu x.M|\!)EC\mathcal{T}(\omega)$.

Consider now the unfoldings of the term $\mu x.(x \oplus M)$. They can be represented as a balanced tree structure, where the nodes are the M^i and row at depth k represents the syntax of $(x \oplus M)^k$, where sibling's nodes are connected by $\oplus$. For instance, we have that

$$(x \oplus M)^0 = M^0,$$
$$(x \oplus M)^1 = M^0 \oplus M^1,$$
$$(x \oplus M)^2 = (M^0 \oplus M^1) \oplus (M^1 \oplus M^2),$$
$$(x \oplus M)^3 = ((M^0 \oplus M^1) \oplus (M^1 \oplus M^2)) \oplus ((M^1 \oplus M^2) \oplus (M^2 \oplus M^3)),\ \text{etc.}$$

For each $\omega \in \Omega$ and each $i \in \mathbb{N}$ there exists $j \leq i$ such that

$$(\!|(x \oplus M)^i|\!)EC\mathcal{T}(\omega) = (\!|M^j|\!)EC(Snd^i(\mathcal{T}))(\omega).$$

In fact, each $\omega \in \Omega$ represents a path in this syntactic tree (e.g., 0 chooses the left branch and 1 chooses the right branch) and the intersection of this path with the i-th level of the graph (representing $(x \oplus M)^i$) is exactly the term M^j satisfying the previous equation.

$((\!|(x \oplus M)^i|\!)EC\mathcal{T}(\omega))_{i \in \mathbb{N}}$ increasingly converges to $(\!|\mu x.M|\!)EC\mathcal{T}(\omega)$, except for those ω for which the sequence stabilises, i.e., for those $\omega \in \Omega$ for which the corresponding path in the syntactic tree always chooses the left branch after a certain level. This is the set

$$S = \{\omega \in \Omega \mid \omega = v0^*, w \in \{0,1\}^*\}.$$

Obviously S is a null set, and this concludes our proof. □

The results we have proven so far allow us to say more about β-reduction. Firstly we prove that an unrestricted version of β-reduction cannot be stated for the stochastic λ-calculus.

Theorem 8.15. *There exist terms M, N (which are not stable) such that*

$$[\![(\lambda x.M)(N) = M\{N/x\}]\!]_{\mathcal{T}} \neq \Omega/_{Null}.$$

Proof. We exploit the results in Theorems 8.5 - 8.3 to derive a contradiction from the assumption that for all terms M, N we have that $[\![(\lambda x.M)(N) = M\{N/x\}]\!]_{\mathcal{T}} = \Omega/_{Null}$.

Consider the following terms

$$\top = \lambda x.\lambda y.x, \quad \bot = \lambda x.\lambda y.y, \quad \mathbf{xor} = \lambda x.\lambda y.x(y\bot\top)(y\top\bot),$$

and let $N = \top \oplus \bot$. Then, consider $M = (\lambda x.\mathbf{xor}\ xx)N$. On the one hand we have that all the following equations have the value $\Omega/_{Null}$ when evaluated for $\mathcal{T}$, due to Theorem 8.1.

$$M = \mathbf{xor}NN = N(N\bot\top)(N\top\bot)$$

Hence, for any closed stable term K,

$$[\![M = K]\!]_{\mathcal{T}} \approx [\![(\top(N\bot\top)(N\top\bot)) \oplus (\bot(N\bot\top)(N\top\bot) = K]\!]_{\mathcal{T}}$$
$$\approx [\![(N\bot\top) \oplus (N\bot\top) = K]\!]_{\mathcal{T}} \approx [\![N\bot\top = K]\!]_{\mathcal{T}}$$
$$[\![(\top\bot\top) \oplus (\bot\bot\top) = K]\!]_{\mathcal{T}} \approx [\![\bot \oplus \top = K]\!]_{\mathcal{T}} \approx [\![\top \oplus \bot = K]\!]_{\mathcal{T}}.$$

On the other hand, we have that for any closed stable term K:

$$[\![M = K]\!]_{\mathcal{T}} \approx [\![(\lambda x.\mathbf{xor}\ xx)\top \oplus (\lambda x.\mathbf{xor}\ xx)\bot = K]\!]_{\mathcal{T}}$$
$$\approx [\![(\mathbf{xor}\top\top) \oplus (\mathbf{xor}\bot\bot) = K]\!]_{\mathcal{T}} \approx [\![\bot \oplus \bot = K]\!]_{\mathcal{T}} \approx [\![\bot = K]\!]_{\mathcal{T}}.$$

Putting together these two sequences of automorphic elements we get that for any closed stable term K,

$$[\![\top \oplus \bot = K]\!]_{\mathcal{T}} \approx [\![\bot = K]\!]_{\mathcal{T}}.$$

Since $\bot$ is a closed stable term, this last equation that

$$[\![\top \oplus \bot = \bot]\!]_{\mathcal{T}} = \Omega/_{Null},$$

meaning $\mathcal{R}(\Omega)$ is a degenerate model, *i.e.* a singleton; this is a contradiction. □

We can, however, have a stronger version of β-reduction than the one stated in Theorem 8.3.

Theorem 8.16 (Extended β-reduction)**.** *If $\lambda x.M_1, \lambda x.M_2$ are closed terms and N is a stable closed term such that for any stable closed term K and any $i \in \{1, 2\}$ we have that*

$$[\![(\lambda x.M_i)N = K]\!]_{\mathcal{T}} \approx [\![M_i\{N/x\} = K]\!]_{\mathcal{T}},$$

then for any stable closed term K,

$$[\![(\lambda x.(M_1 \oplus M_2))N = K]\!]_{\mathcal{T}} \approx [\![M_1\{N/x\} \oplus M_2\{N/x\} = K]\!]_{\mathcal{T}}.$$

Proof. Note that

$$(\!|(\lambda x.(M_1 \oplus M_2))N|\!)\emptyset id\mathcal{T}$$
$$= (\!|\lambda x.(M_1 \oplus M_2)|\!)\emptyset(\lambda a.(\!|N|\!)\emptyset(\lambda b.id(\phi(a)b))\mathcal{T}^e)T^o$$

and applying λ-distributivity,

$$= (\!|\lambda x.M_1 \oplus \lambda x.M_2)|\!)\emptyset(\lambda a.(\!|N|\!)\emptyset(\lambda b.id(\phi(a)b))\mathcal{T}^e)T^o$$

now we solve the probabilistic choice and get some $j \in \{1, 2\}$ such that

$$= (\!|\lambda x.M_j|\!)\emptyset(\lambda a.(\!|N|\!)\emptyset(\lambda b.id(\phi(a)b))\mathcal{T}^e)\mathbf{Snd}(T^o)$$
$$(\!|(\lambda x.M_j)N|\!)\emptyset id\mathcal{T}',$$

where $\mathcal{T}' = \mathbf{Swap}(\mathbf{Snd}(\mathcal{T}))$. Hence, we have that

$$[\![(\lambda x.(M_1 \oplus M_2))N = K]\!]_{\mathcal{T}}$$
$$= \{\omega \in \Omega \mid (\!|(\lambda x.M_1 \oplus \lambda x.M_2))N|\!)\emptyset id\mathcal{T}(\omega) = (\!|K|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null}$$
$$= \{\omega \in \Omega \mid (\!|(\lambda x.M_j)N|\!)\emptyset id\mathcal{T}'(\omega) = (\!|K|\!)\emptyset id\mathcal{T}(\omega)\}/_{Null}.$$

Since K is stable, $(\!|K|\!)\emptyset id\mathcal{T} = (\!|K|\!)\emptyset id\mathcal{T}'$, so

$$= \{\omega \in \Omega \mid (\!|(\lambda x.M_j)N|\!)\emptyset id\mathcal{T}'(\omega) = (\!|K|\!)\emptyset id\mathcal{T}'(\omega)\}/_{Null}$$
$$= [\![(\lambda x.M_j)N = K]\!]_{\mathcal{T}'} \approx [\![M_j\{N/x\} = K]\!]_{\mathcal{T}'}.$$

Similarly we get that

$$[\![M_1\{N/x\} \oplus M_2\{N/x\} = K]\!]_{\mathcal{T}} \approx [\![M_j\{N/x\} = K]\!]_{\mathcal{T}''},$$

for some tossing $\mathcal{T}''$. Further, Theorema 7.2 ensures us that

$$[\![M_j\{N/x\} = K]\!]_{\mathcal{T}'} \approx [\![M_j\{N/x\} = K]\!]_{\mathcal{T}''},$$

which concludes our proof, since $\approx$ is transitive. □

9 Generating Random Numbers

In this section we present a small example of programming in stochastic λ-calculus and use the semantics to argue for the correctness of the program behavior. Our program takes a Church numeral n and produces a random Church numeral from 0 to $2^n - 1$ with equal probability.

Functions for Church Numerals and Booleans. In the following, we use mathematical symbols as the names of lambda terms, for ease of notation. Note that this means arithmetical expressions are in (forward) Polish notation.

Recall that a Church numeral for the number $n \in \mathbb{N}$ is a function of two arguments f and x, returning f applied n times to x. We use the well-known encodings of the arithmetic operations succ, + and ×.

Picking a Random Number from 0 to $2^n - 1$. The following stochastic λ-term is the key of our encoding.

$$\mathtt{rand} = \lambda n.n(\lambda x.((\times\ 2\ x) \oplus (\mathsf{succ}(\times\ 2\ x))))\emptyset.$$

In intuitive terms, the program starts with a number equal to 0 and flips a fair coin n times, either doubling the number or doubling and adding one, depending on the outcome.

The following statement, if demonstrated, proves that the program has the desired behavior. It exemplifies how our deduction principles can be applied.

Statement: For all Church numerals n and tossing processes $\mathcal{T}$,

$$\bigvee_{i=0}^{2^n-1} [\![\mathtt{rand}\, n = i]\!]_{\mathcal{T}} = \Omega/_{Null}$$

and for all $0 \le i, j \le 2^n - 1$

$$[\![\mathtt{rand}\, n = i]\!]_{\mathcal{T}} \approx [\![\mathtt{rand}\, n = j]\!]_{\mathcal{T}}.$$

Therefore, for each $0 \le i \le 2^n - 1$, $[\![\mathtt{rand}\ n = i]\!]_{\mathcal{T}}$ are of equal probabilities and summing to 1, *i.e.* of probability 2^{-n}.

Sketch: We do not provide a detailed proof of this statement, that would require further developments of the deduction principles. Instead, we sketch below how such a proof shall be organized.
Induction on n: the inductive hypothesis we need is that
$\bigvee_{i=0}^{\infty} [\![\mathtt{rand}\, n = i]\!]_{\mathcal{T}} = \Omega/_{Null}$, that for all $i \ge 2^n$, we have
$[\![\mathtt{rand}\, n = i]\!]_{\mathcal{T}} = \emptyset/_{Null}$, and for all $0 \le i, j \le 2^n - 1$, we have
$[\![\mathtt{rand}\, n = i]\!]_{\mathcal{T}} \equiv [\![\mathtt{rand}\, n = j]\!]_{\mathcal{T}}$.

We start with the base case, $n = 0$. Then

$$\mathtt{rand}\ \mathtt{0} = \mathtt{0}(\lambda x.((\times\ 2\ x) \oplus (\mathsf{succ}(\times\ 2\ x))))\mathtt{0} = \mathtt{0}.$$

We have $[\![\mathtt{rand}\, 0 = 0]\!] = \Omega/_{Null}$. This shows the three facts we want, because $0 \le i \le 2^0 - 1 = 0$ implies $i = 0$.

For the inductive step, we temporarily introduce the name

$$f' = (\lambda x.((\times\ 2\ x) \oplus (\mathsf{succ}(\times\ 2\ x))))\,,$$

for the latter part of the definition of rand, excluding the final 0. We start by re-expressing $\mathsf{rand}\,(\mathsf{succ}\ n)$ in terms of $\mathsf{rand}\ n$.

$$\begin{aligned}\mathsf{rand}(\mathsf{succ}\ n) &= (\mathsf{succ}\ n)f'\mathtt{0} = f'(nf'\mathtt{0}) = f'(\mathsf{rand}\ n)\\ &= (\lambda x.((\times\ 2\ x) \oplus (\mathsf{succ}(\times\ 2\ x))))(\mathsf{rand}\ n)\\ &= ((\lambda x.\times\ 2\ x) \oplus (\lambda x.\mathsf{succ}(\times\ 2\ x)))(\mathsf{rand}\ n)\\ &= ((\lambda x.\times\ 2\ x)(\mathsf{rand}\ n)) \oplus ((\lambda x.\mathsf{succ}(\times\ 2\ x))(\mathsf{rand}\ n))\,.\end{aligned}$$

By the inductive hypothesis, we have $\bigvee_{i=0}^{\infty} [\![\mathsf{rand}\ n = i]\!]_{\mathcal{T}} = \Omega/_{Null}$. By well-definedness of function application, it holds that

$$\bigvee_{i=0}^{\infty} [\![(\lambda x.\times\ 2\ x)(\mathsf{rand}\ n) = i]\!]_{\mathcal{T}} = \Omega/_{Null}\,,$$

$$\bigvee_{i=0}^{\infty} [\![(\lambda x.\mathsf{succ}(\times\ 2\ x))(\mathsf{rand}\ n) = i]\!]_{\mathcal{T}} = \Omega/_{Null}\,.$$

so, because of the properties of $\oplus$ (Theorems 8.5 and 8.10),

$$\bigvee_{i=0}^{\infty} [\![\mathsf{rand}\,(\mathsf{succ}\ n) = i]\!]_{\mathcal{T}} = \Omega/_{Null}.$$

For the second part of the inductive hypothesis, that $i \geq 2^{n+1}$ implies $[\![\mathsf{rand}\,(\mathsf{succ}\ n) = i]\!]_{\mathcal{T}} = \emptyset/_{Null}$, we make a case split depending on whether i is even or odd. As both cases are similar, we only show the case where i is even. Then $i = 2i'$ for some integer i'. By the inductive hypothesis $[\![\mathsf{rand}\ n\ =\ i']\!]_{\mathcal{T}} = \emptyset/_{Null}$, so by well-definedness and the fact that $(\times 2i') = i$, we have $[\![\mathsf{rand}\,(\mathsf{succ}\ n) = i]\!]_{\mathcal{T}} = \emptyset/_{Null}$.

Finally, we need to show that for all $0 \leq i, j \leq 2^{n+1} - 1$ we have $[\![\mathsf{rand}\,(\mathsf{succ}\ n) = i]\!]_{\mathcal{T}} \approx [\![\mathsf{rand}\,(\mathsf{succ}\ n) = j]\!]_{\mathcal{T}}$. We have a four-way case split according to whether i and j are odd or even. We treat the case where i and j are even, as the other three cases are similar. Let i', j' such that $i = 2i'$ and $j = 2j'$. Hence,

$$\begin{aligned}&[\![\mathsf{rand}(\mathsf{succ}\ n) = i]\!]_{\mathcal{T}}\\ &= [\![(\lambda x.\times 2\,x)(\mathsf{rand}n) \oplus ((\lambda x.\mathsf{succ}(\times 2\,x))(\mathsf{rand}n)) = (\lambda x.\times 2\,x)i']\!]_{\mathcal{T}}\\ &\approx [\![(\lambda x.\times\ 2\ x)(\mathsf{rand}\ n) = (\lambda x.\times\ 2\ x)i']\!]_{\mathcal{T}}\\ &= [\![\mathsf{rand}\ n = i']\!]_{\mathcal{T}} = [\![\mathsf{rand}\ n = j']\!]_{\mathcal{T}} = [\![\mathsf{rand}(\mathsf{succ}\ n) = j]\!]_{\mathcal{T}}\,.\end{aligned}$$

10 Conclusions

We see this paper as the beginning of an investigation into random processes at higher type. There are many things to investigate:

- We need to understand how this relates to more categorical approaches [7] based on the cartesian closed category of quasi-Borel spaces.
- We need to develop a deeper understanding of the Boolean-valued reasoning principles that we have used here.
- It would be very interesting to develop suitable dependently-typed versions of a stochastic λ-calculus; indeed this was one of the main motivations of [18].
- The relation between invariance results as we have used them and exchangability and symmetry principles in probability theory (see, for example, [10]) need to be understood better.

Acknowledgments

This research was supported by the DFF Danish research grant FNU 4181-00360, by a grant from NSERC (Canada), and by a grant from the National Science Foundation (USA). We gratefully acknowledge the support of the Simons Institute Logical Structures in Computation Program in Fall 2016. We thank Ugo Dal Lago, Cameron Freer, Marco Gaobardi, Chris Heunen, Alex Simpson and Sam Staton for useful discussions.

References

[1] John L. Bell. 1985. *Boolean-Valued Models and Independence Proofs in Set Theory* (2nd ed.). Number 12 in Oxford Logic Guides. Oxford University Press.
[2] David H. Fremlin. 2003. Measure Theory, Volume 4. http://www.essex.ac.uk/maths/people/fremlin/mt.htm. (2003).
[3] R. M. Friedberg and H. Rogers. 1959. Reducibility and Completeness for sets of integers. *Mathematical Logic Quarterly* 5 (1959), 117–125.
[4] Noah Goodman, Vikash Mansingkha, Daniel Roy, Keith Bonawitz, and Joshua Tenenbaum. 2008. Church: a language for generative models. In *Proceedings of the 24th Conference on Uncertainty in Artificial Intelligence*. 220–229.
[5] Andrew D. Gordon, Thore Graepel, Nicolas Rolland, Claudio V. Russo, Johannes Borgström, and John Guiver. 2014. Tabular: a schema-driven probabilistic programming language. In *Proceedings of POPL '14, San Diego, CA, USA, January 20-21, 2014*. 321–334.
[6] V. Gupta, R. Jagadeesan, and P. Panangaden. 1999. Stochastic Processes as Concurrent Constraint Programs. In *Proceedings of the 26th Proceedings Of The Annual ACM Symposium On Principles Of Programming Languages*. 189–202.
[7] Chris Heunen, Ohad Kammar, Sam Staton, and Hongseok Yang. 2017. A convenient category for higher-order probability theory. In *Proceedings of the Thirty-second Annual ACM-IEEE Symposium on Logic in Computer Science*.
[8] C. Jones and G. D. Plotkin. 1989. A Probabilistic Powerdomain of Evaluations. In *Proceedings of the Fourth Annual IEEE Symposium On Logic In Computer Science*. 186–195.
[9] A. Jung and R. Tix. 1998. The troublesome probabilistic powerdomain. *Electronic Notes in Theoretical Computer Science* 13 (1998).
[10] Olav Kallenberg. 2006. *Probabilistic symmetries and invariance principles*. Springer Science and Business Media.
[11] D. Kozen. 1981. Semantics of Probabilistic Programs. *Journal of Computer and Systems Sciences* 22 (1981), 328–350.
[12] Ugo Dal Lago, Davide Sangiorgi, and Michele Alberti. 2014. On Coinductive Equivalences for Higher-order Probabilistic Functional Programs. In *Proceedings of the 41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL '14)*. ACM, New York, NY, USA, 297–308.
[13] Norman Ramsey and Avi Pfeffer. 2002. Stochastic lambda calculus and monads of probability distributions. In *The 29th SIGPLAN-SIGACT Symposium on Principles of Programming Languages*. 154–165.
[14] H. Rogers. 1967. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill.
[15] Vladimir A. Rohlin. 1952. On the Fundamental Ideas of Measure Theory. *Amer. Math. Soc. Transl.* 71 (1952).
[16] N. Saheb-Djahromi. 1978. Probabilistic LCF. In *Mathematical Foundations Of Computer Science (Lecture Notes In Computer Science)*. Springer-Verlag.
[17] Dana Scott. 1967. A Proof of the Independence of the Continuum Hypothesis. *Mathematical Systems Theory* 1, 2 (1967), 89–111.
[18] D. Scott. 2014. Stochastic Lambda Calculi: An extended Abstract. *Journal of Applied Logic* 12 (2014), 369–376.
[19] S. Smolka, P. Kumar, N. Fosterand D. Kozen, and A. Silva. 2017. Cantor meets Scott: Domain-theoretic foundations for probabilistic network programming. In *Proceedings of the 44th ACM SIGPLAN Symp. Principles of Programming Languages (POPL'17)*. ACM, 557–571. ACM SIGPLAN Notices - POPL '17, Volume 52 Issue 1.
[20] Sam Staton, Hongseok Yang, Frank Wood, Chris Heunen, and Ohad Kammar. 2017. Semantics for probabilistic programming: higher-order functions, continuous distributions, and soft constraints. In *Proceedings of the 31st Annual ACM-IEEE Symposium On Logic In Computer Science*. 525–534.
[21] David Tolpin, Jan-Willem van de Meent, Hongseok Yang, and Frank Wood. 2016. Design and implementation of probabilistic programming language anglican. In *Proceedings of the 28th Symposium on the Implementation and Application of Functional Programming Languages*. ACM, 6.
[22] David Williams. 1991. *Probability with Martingales*. Cambridge University Press.